EIGENGAZE - COVERT BEHAVIORAL BIOMETRIC EXPLOITING VISUAL ATTENTION CHARACTERISTICS

Clinton Fookes and Sridha Sridharan

Image & Video Research Laboratory, Queensland University of Technology
2 George St GPO Box 2434 Brisbane QLD 4001 Australia
{c.fookes,s.sridharan}@qut.edu.au

ABSTRACT

It is possible for the visual attention characteristics of a person to be exploited as a biometric for authentication or identification of individual viewers. The visual attention characteristics of a person can be easily monitored by tracking the gaze of a viewer during the presentation of a known or unknown visual scene. The positions and sequences of gaze locations during viewing may be determined by overt (conscious) or covert (sub-conscious) viewing behaviour. This paper presents a method to authenticate individuals using their covert viewing behaviour, thus yielding a unique behavioural biometric. A method to quantify the spatial and temporal patterns established by the viewer for their covert behaviour is proposed utilsing a principal component analysis technique called 'eigenGaze'. Experimental results suggest that it is possible to capture the unique visual attention characteristics of a person to provide a simple behavioural biometric.

1. INTRODUCTION

Biometrics are increasingly being utilised for their automatic authentication ability in a diverse range of applications including access control, border control, finance, healthcare, law enforcement, transportation, and even entertainment. Biometrics provide the ability to identify or authenticate an individual with much stronger certainty than with knowledge-based or token-based systems alone as they rely on a person's physiological or behavioural traits [1]. However, biometrics do not provide a perfect solution. It is now widely acknowledged that most biometrics are not secret; they are often freely available and easy to acquire, making many current biometric systems prone to spoof attacks (for example, a photo of a person's face, or a copy of their fingerprint can be easily acquired and utilised to foil a biometric system).

While biometrics are harder to violate than pure knowledge-based or token-based systems, it is not impossible. Ideally, safer forms of biometrics would be based on non-visible and non-physiological information hidden deep within the person. Such a biometric could be based on behaviour or even thought processes. Gait recognition is an example of a behavioural biometric which is particularly hard to reproduce even though it is still visible [2]. The downside for such biometrics is generally an increase in the intra-class variability, making the recognition process significantly harder and thus reducing recognition accuracies.

This paper proposes to exploit the personal aspects of an individual's visual attention processes as a unique behavioural biometric which can be used for authentication or identification applications. The visual attention process of a person can be easily inferred through capturing their gaze (i.e. monitoring their viewing behaviour through measuring their eye movements). The approach relies on the principle that the human visual system requires the eye to rest motionless for short periods during its traversal of the viewing space, to assimilate detail at a given location in a visual scene. The positions and sequences of gaze locations during viewing may be determined by overt (conscious) or covert (sub-conscious) viewing behaviour. Overt viewing behaviour can be captured through a simple biometric which mimics the process of entering a PIN number but through a person's gaze pattern [3]. Covert veiwing behaviour, however, is much more problematic to manage. The assumption behind such a biometric is that the viewing behaviour a person employs to gather information from a visual scene is inherently unique to that person. By detecting how and when the viewer looks at certain features in a presented scene, a signature for personal identification can be established and this is completely independent of other physiological measures of the viewer.

The advantages of a biometric based on an individual's visual attention characteristics are numerous. Firstly, such an approach does not require physical contact between the viewer and a device as it is possible to acquire the necessary information on a person's gaze from short to medium distances. Secondly, such a biometric is difficult to detect by third person observation or surveillance due to the ambiguity of the visual attention process. Finally, largely due to the previous reason, such a biometric would be virtually impossible to spoof as the biometric measure itself is comprised of a complex network of characteristics including: the scene under investigation; what the person examined and the order in which they examined; and how their brain processes visual information. This process is heavily related to the higher cognitive, psychological and neurological processes of a person and thus provides the opportunity for a close-to-ideal biometric. The development of a person's subconscious viewing behaviour also begins early in life during infancy stages [4], thus providing further evidence to the uniqueness of this process and giving further weight to its use as a biometric. The ultimate challenge then becomes one of developing a robust biometric system that can identify and extract salient features which hopefully provide the ideal situation where intra-class variability is small and separation between classes (or inter-class variability) is large.

The outline of this paper is as follows. Section 2 will provide some background on the human visual system. Section 3 will present an 'eigenGaze' technique as a behavioural biometric based on an individual's covert viewing behaviour. Section 4 will present some experimental results conducted on a small test database which suggests the feasibility of such a biometric. Finally, the paper is concluded in Section 5.

2. VISUAL ATTENTION PRELIMINARIES

Visual attention is the name of the research field which investigates aspects of human vision and how it relates to higher cognitive, psychological and neurological processes. The concept of attention, or conscious selecting and directing of perceptual information intake, arises because finite physical human limitations prevent us from perceiving all things at once. This process makes efficient the serial searching and processing of areas for "visual processing" by using the scan paths of the eyes to simplify an image to extract relevant information based upon the task at hand [5].

The human visual system relies on positioning of the eyes to bring a particular component of the visible field of view into high resolution. This permits the person to view an object or region of interest near the centre of the field in much finer detail. In this respect, visual attention acts as a "spotlight" effect [6]. The region viewed at high resolution is known as the foveal region and is much smaller than the entire field of view contained in the periphery. Viewing of a visual scene consists of a sequence of brief gaze periods (typically 100-500ms) of visual concentration (fixations) at specic locations in the field of view, interspersed with sudden movements of the eyes (saccades) to reposition the foveal region at the next point of attention. This process provides the brain with detailed visual information over a succession of these fixation-saccade events covering a few comparatively small areas in the field of view. From this process, a "conceptual" image of the visual scene is constructed by combining these fixation-saccade events with the large area of low resolution information gained from the periphery. The fixation-saccade events may be consciously directed by the viewer to visit a sequence of specic points in the scene (overt), or else may be allowed to be directed sub-consciously by the brain according to its choice of points of interest (covert) [7]. Thus, by observing where and when a person's gaze is directed, it is possible to establish the fixation-saccade path followed by the viewer. This provides insights about what the viewer found interesting (i.e. what captured their attention) and perhaps reveal how that person perceived the visual scene they were viewing.

3. 'EIGENGAZE' TECHNIQUE

One of the most common feature extraction techniques employed within the field of biometrics is based on Principal Component Analysis (PCA). This technique was first utilised in a fully automated face recognition system proposed by Turk and Pentland [8]. This work applied PCA to derive a set of face representations which were termed 'eigenFaces'. This technique has since become the gold standard for face verification and was utilised as the baseline system in the Face Recognition Grand Challenge evaluation [9]. Since its introduction, similar techniques have been proposed for other biometric applications including 'eigenLips' for speech recognition, 'eigenEyes' [10], and 'eigenPulse' for human identification from cardiovascular function [11]. This paper proposes a similar 'eigenGaze' technique as a method for developing a biometric from a human's gaze behaviour.

This technique applies eigen-decomposition to the covariance matrix of a set of M vectorised training sequences of gaze. PCA is used to derive a set of eigenvectors which are ranked based on their eigenvalues λ. The D most relevant eigenvectors are retained to form a sub-space ϕ. The eigenvalues represent the variance of each eigenvector and so represent the relative importance of each of the eigenvectors with regards to minimising the reconstruction error in a least squares sense. Once the sub-space ϕ is obtained, a vectorised gaze sequence $\mathbf{v}_a$ can be projected into the space to obtain a feature vector, **a**,

$$\mathbf{a} = (\mathbf{v}_a - \omega) \times \phi, \tag{1}$$

where ω is the mean gaze vector. This technique can be termed 'eigenGaze' as each eigenvector is representative of the most variant attributes of the training gaze sequences (similar to eigenFaces for face recognition). As the eye movements of a viewer were recorded during a viewing process, the gaze data was passed through a clustering algorithm to extract fixations with a time constraint (or threshold) imposed on the fixation duration, T_{fix} [12].

After the gaze sequence was clustered into fixations, the mean duration of clusters, the mean number of revists, cluster length and mean fixation durations, were then extracted from the data and used as features for classification. To find the eigenGazes, each gaze capture is converted into a vector of clustered fixations, Γn, of length 20. Multiple gazes per person are utilised as this sharply increases accuracy due to the increased information available on each known individual. This collection of gazes can be referred to as the gaze space.

As usual for the construction of a basis set, the mean of the observations is removed and a covariance matrix, C, for the dataset is computed. The eigenGazes then are simply the eigenvectors of C. These eigenGazes provide a small yet powerful basis for the gaze space. Using only a weighted sum of these eigenGazes, it is possible to reconstruct each gaze in the dataset.

4. EXPERIMENTAL RESULTS

The device used to record eye movements during the experiments was an EyeTech video-based corneal reflection eye tracker. This device is normally used for point-of-regard measurements, i.e. those that measure the position of the eye relative to 3D space rather than relative to the head. The method of operation relies on tracking the corneal reflection from an infra-red light source, as it is invisible to the human eye and non-distracting. Although four separate reflections are formed during the reflection of a light source, referred to as Purkinje reflections, only the first Purkinje reflection is utilised in the video-based tracker and is located relative to the location of the pupil centre using image processing techniques. A small amount of head movement is acceptable for stable gaze monitoring with this device. The gaze-tracker utilised operated in a default setting of 15 frames per second, resulting in a sample of the observer's gaze direction approximately every 67ms. The experiments were conducted using an image and screen resolution of 1024×768 pixels.

The experimental methodology adopted consisted of recording gaze data of five different viewers for a particular image of an outdoor scene of a rock climb (see Figure 1). For each session, the viewer was directed to examine the scene without any consciously directed (or task specified) gaze pattern. That is, the viewer was free to examine the image in their natural manner, i.e. at the whim of their personal visual attention processes. The sequence of eye saccades and fixations was captured for a total duration of 10 secs. For each case, the gaze-tracking experiment was repeated three times, each occasion being separated from the others by several days or by some other visual tasks to reduce the influence of repetition.

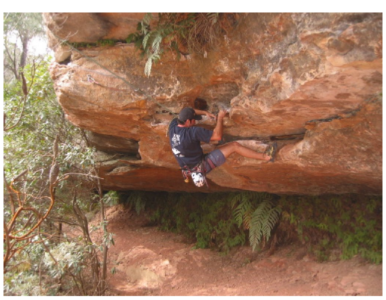

Fig. 1. Rockclimb image used in the covert experiments.

Figure 2 presents some sample data for the experiments. Plots (a), (b) and (c) show fixations of Scan 1 for Person 1, Person 2, and Person 3 (out of the database of 5 people) respectively, plotted against the Rockclimb image. The variations between observers in these cases are quite apparent. Figures (d), (e) and (f) represent the three repeated scans for Person 1 on the Rockclimb image. These three scans appear to have many similarities after a visual comparison, however, there is still obviously some diversity between them, i.e. intra-viewer variation. Moreover, the plots in Figure 2 do not contain any information about the sequence in which these points were viewed.

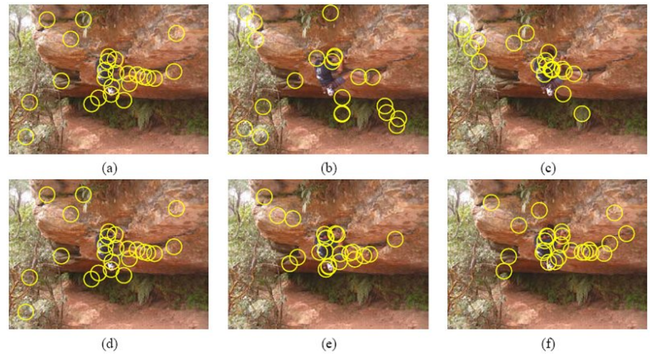

Fig. 2. Covert Gaze Data: Extracted fixations for Scan 1 of (a) Person 1, (b) Person 2, and (c) Person 3, for Rockclimb image. Figures (d), (e) and (f) show all extracted fixations from the three repeated scans for Person 1.

For each view of the Rockclimb image, the first eight eigenGazes (which comprised 99.5% of the gaze space) were used to extract a feature vector of weights which were then passed to the classifier to evaluate the probability of the gaze belonging to a given individual. The similarity measure used in this case is the cosine of the angle between the projected eigenGaze vectors. Table 1 presents the similarity measures calculated between all possible scan combinations (where P_i is the ith person being tested, and V_j is the jth viewing sequence for a given person).

As illustrated in Table 1, the intra-class similarity scores (bolded) are generally much higher (close to unity) than the inter-class scores. Based on this dataset it is clear that there is strong separation between classes which should yield reasonable classification rates. That is, intra-class scores of a person's gaze viewing sequence matching to another of their scans at a later period in time is higher than the score achieved from matching with any of the other four identities in the database.

To illustrate the recognition ability of the 'eigenGaze' method, the scores in Table 1 were utilised to generate a Detection Error Trade-off plot (or DET plot). This is given in Figure 3 which shows the false alarm probability versus the miss probability for the 'eigenGaze' technique. This method resulted in an equal error rate of 8.9% as shown by the black circle plotted in Figure 3.

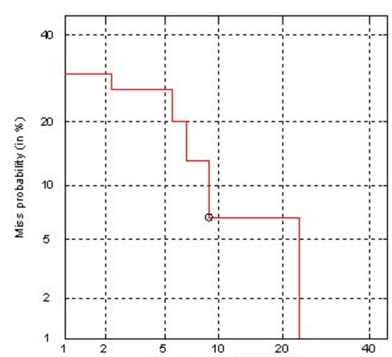

Fig. 3. Detection-Error-Trade-off plot for the eigenGaze covert experiment yielding an equal error rate of 8.9%.

	P1 V1	P1 V2	P1 V3	P2 V1	P2 V2	P2 V3	P3 V1	P3 V2	P3 V3	P4 V1	P4 V2	P4 V3	P5 V1	P5 V2	P5 V3
P1V1	**1**	**0.99**	**0.85**	-0.1	-0.0	0.3	-0.4	-0.4	-0.4	-0.4	-0.3	-0.4	-0.9	-0.4	-0.9
P1V2		**1**	**0.81**	-0.2	-0.1	0.2	-0.4	-0.4	-0.4	-0.5	-0.3	-0.4	-0.9	-0.5	-0.9
P1V3			**1**	0.4	0.4	0.7	-0.6	-0.6	-0.6	-0.5	-0.5	-0.5	-0.7	0.12	-0.7
P2V1				**1**	**1**	**0.93**	-0.2	-0.1	-0.1	0.11	-0.1	0.1	0.32	0.92	0.34
P2V2					**1**	**0.93**	-0.2	-0.2	-0.2	0.08	-0.1	0.1	0.30	0.92	0.33
P2V3						**1**	-0.2	-0.2	-0.1	0.09	-0.0	0.1	0.06	0.71	0.07
P3V1							**1**	**1**	**1**	0.97	0.99	0.96	0.66	-0.1	0.59
P3V2								**1**	**1**	0.46	0.45	0.45	0.67	-0.1	0.61
P3V3									**1**	0.97	0.99	0.97	0.68	-0.1	0.62
P4V1										**1**	**0.97**	**1**	0.75	0.12	0.70
P4V2											**1**	**0.97**	0.6	-0.1	0.53
P4V3												**1**	0.73	0.11	0.67
P5V1													**1**	**0.54**	**1**
P5V2														**1**	**0.59**
P5V3															**1**

Table 1. Classification results generated by 5 people with 3 different sequences projected into the eigenGaze space.

This experiment does yield promising results that the 'eigenGaze' approach may be a worthy technique capable of classifying individuals using the clustering of their gaze data. Although this technique does certainly yield some potential, the true capability of this technique and the viability of eigenGaze will not be clear until further research and experiments are conducted on much larger datasets.

5. CONCLUSION AND FUTURE WORK

This paper has presented how the use of a person's covert visual attention characteristics can be employed as a behavioural biometric for the authentication or identification of an individual. The visual attention characteristics of a person can be easily monitored by observing where and when a person's gaze is directed. Establishing this fixation-saccade path followed by the viewer through gaze tracking provides strong insights into a persons internal visual attention process through observing what captured their attention and how they perceive a visual scene. A method to quantify the spatial and temporal patterns of the covert gaze behaviours was proposed through an 'eigenGaze' technique involving a Principal Component Analysis or eigen-decomposition approach which was applied to clusters of gaze positions. Experimental results suggest the proposed technique can provide a simple and effective biometric for classifying individuals. The method yielded an equal error rate of 8.9% on a small database of viewers. Further research is required to assess and confirm the true capability and behaviour of this proposed gaze biometric through testing and analysis of the discriminability of gaze features across databases of a larger number of viewers and images to provide a richer set of base data. Future opportunities for such a gaze biometric to be utilised in liveness detection and anti-spoofing measures within biometric systems is also very promising.

6. REFERENCES

[1] N. Ratha, J. Connell, and R. Bolle, "Biometrics break-ins and band-aids," *Pattern Recognition Letters*, vol. 24, pp. 2105 2113, 2003.

[2] M.S. Nixon and J.N. Carter, "Automatic recognition by gait," *Proceedings of the IEEE*, vol. 94, no. 11, pp. 2013 – 2024, 2006.

[3] A. Maeder, C. Fookes, and S. Sridharan, "Gaze based user authentication for personal computer applications," in *In Proceedings of the 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, Hong Kong*, 2004, p. 727 730.

[4] A. Belardinelli, F. Pirri, and A. Carbone, "Bottom-up gaze shifts and fixations learning by imitation," *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 37, no. 2, pp. 256 271, 2007.

[5] D. Noton and L. Stark, "Eye movements and visual perception," *Scientific American*, vol. 224, no. 6, pp. 35–43, 1971.

[6] L. Itti and C. Koch, "Feature combination strategies for saliency-based visual attention systems," *Journal of Elect. Imag.*, vol. 10, no. 1, pp. 161–169, 2001.

[7] L. Itti and C. Koch, "A saliency-based search mechanism for overt and covert shifts of visual attention," *Vision Research*, vol. 40, pp. 1489 1506, 2000.

[8] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[9] J. Phillips, P. Flynn, T. Scruggs, K. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in *IEEE Conference of Computer Vision and Pattern Recognition*, 2005, vol. 1, pp. 947–954.

[10] T.E. Campos, R. Feris, and R.M. Cesar Jr, "Eigenfaces versus eigeneyes: First steps toward performance assessment of representarions for face recognition," *Lecture Notes in Artificial Intelligence*, vol. 1793, pp. 197–206, 2000.

[11] J.M. Irvine, S.A. Israel, W.T. Scruggs, and W.J. Worek, "eigenpulse: Robust human identification from cardiovascular function," *Pattern Recognition*, vol. 41, no. 11, pp. 3427 3435, 2008.

[12] A. Maeder and C. Fookes, "A visual attention approach to personal identification," in *In Proceedings of the Eighth Australian and New Zealand Intelligent Information Systems Conference*, 2003, pp. 55–60.